Supplementary Table 3| Round 2 Questions to the NEWG

Topic/Rating Instructions	Statements
Routine Assessment Please rate the extent to which you agree or disagree with the importance of each of the following factors in selecting particular medications for routine human research:	• Central Nervous System acting medications. • Medications with documented impact on the physical development of the fetus (i.e. malformation, dysmorphic facial features, growth restriction). • Evidence of risk from animal studies. • Medications that are widely used in female populations and thus the rate of fetal exposure is high. • Is related either by therapeutic class or mechanism of action to a medication known already to have impact on the development of the human brain • Published case reports, pilot studies or other early evidence indicating potential risk or association with impairment. • A signal or risk generated through industry or regulatory adverse event reporting systems.
Pre-Clinical Studies	• Where animal studies indicate risk for a specific medication this medication should automatically be subject to routine follow-up in humans. • Where animal studies do NOT indicate risk for a specific medication this cannot be interpreted as evidence of safety in humans.
Barriers to Research Please indicate the extent to which you disagree or agree that each of the following represents a barrier to research in this area: Please rate your level of agreement with the following statement:	• Difficulty obtaining funding is a barrier. • A view that neurodevelopmental functioning is less important or of less significance than physical child outcomes is a barrier. • Cost of follow-up to investigate long-term effects is a barrier • An international lack of expertise required for assessment and interpretation of neurodevelopmental outcomes is a barrier • Lack of consensus regarding approach to study design is a barrier • In an ideal world, in which funding is ample, many of the barriers to child neurodevelopmental research can be overcome by using optimal study design
Single vs Multiple Scores/Domains	• To what extent do you agree that no single score or domain can reliably inform on the functioning of all other neurodevelopmental domains?

Core Outcomes
 	
Please indicate your level of agreement or disagreement that each of the domains listed below can be altered by prenatal exposure.
	• Cognitive
	  o General cognitive skills (e.g. developmental quotient, intelligence quotient, IQ); Expressive Language; Receptive Language; Executive Function; Learning rate; Memory; Attention; Processing; Visuo-Spatial skills
	• Social/Behavioural
	  o Adaptive behaviour (e.g. daily living skills); Behavior Problems; Social skills
	• Motor
	  o Developmental milestones; Reflexes; Fine motor; Gross motor
	• Emotion and mood difficulties
	  o Emotional regulation difficulties; Anxiety; Depression
	• 'Real world' outcomes
	  o Examination results; Rates of specialist educational needs/ support; Occupational attainment
	• Clinical Disorders

Core Domains	• Cognitive

Please indicate what you think should be included in a **CORE** or **CENTRAL** set of neurodevelopmental outcomes that would require investigation across studies before conclusions can be made regarding the potential longer term impact on child neurodevelopment.	• Adaptive behaviour (e.g. daily living skills) • Social skills • Motor • Behavior Problems • Emotion and mood difficulties • 'Real world' outcomes • Clinical Disorders • Sexual Function
Optimal Study Design Please indicate your level of agreement or disagreement with the following statement:	• A variety of investigative approaches including prospective, longitudinal studies, routine healthcare records, and animal data is required to ensure optimisation of data collection in this area.
Optimal Design Features Please indicate your agreement or disagreement with the following statements regarding an optimal approach to outcome measurement:	• The optimal research strategy for neurodevelopmental domains would involve direct assessment of the child for the purpose of the study? • For some neurodevelopmental skills or domains such as Cognitive or Motor functioning researcher administered assessments would be optimum, whilst for others domains parental completed questionnaires may be more applicable (e.g. behaviour, social skills, mood). • Measurements of the different aspects of neurodevelopment (whether researcher or parent completed) should be undertaken using standardised and validated measures? • Assessments should be administered by experts in the relevant field.
Optimal Timing Please indicate your level of agreement or disagreement with the following statements:	• The optimal research design for investigations of neurodevelopmental outcomes would involve long-term follow-up that begins in early infancy and continues into early adulthood. • Many neurodevelopmental impairments may only become fully apparent as the child gets older. • Assessments in infancy (=/<24 months) are useful for identifying early deviations in neurodevelopmental development. • Assessments in school-age children (4-5+ years) provide more useful and reliable information than assessments carried out during infancy (0-2 years). • Assessments in the early adulthood would provide the most reliable information regarding functioning as all neurodevelopmental domains are reaching full maturity.
Please indicate what you think is the ideal starting point for neurodevelopmental assessments following exposure to medication in the womb:	• In utero • 6 months • 2- years • 4-5 years
Please indicate what you think is the ideal (funding no issue) end-point for neurodevelopmental assessments following exposure to medication in the womb:	• 4-5 years • 8-10 years • 12 years • 16 years • 21 years • 30 years
Can final conclusions regarding the level of risk a medication is associated with be drawn following assessments in infancy only?	• Yes • No
Optimal Intervals	• It is important to develop an assessment protocol across different stages of development in order to facilitate the building of a

Please indicate your agreement or disagreement with the following statements: Please indicate your agreement or disagreement with the following statements:	comprehensive understanding of the developmental trajectory of core neurodevelopmental outcomes • Assessments should be more frequent during infancy and decrease as the child gets older. • Development occurs at a much higher rate during infancy and so requires more regular assessments. • Development occurs at a slower rate once the child reaches school age and so less regular assessments are required. • Early deviation in neurodevelopmental trajectories following exposure to medication in utero suggest that more intensive follow-up is required. • Lack of early deviation in neurodevelopmental trajectories does not mean later effects will not emerge, and so regular follow-up is still required.

Confounding
 	
Tick all that you feel should be included in an optimal study design when the primary outcome is a neurodevelopmental domain:
	• Maternal Factors
	  o Indication; IQ; Education; Age; Parity; Marital Status; Socio-economic status; Nutrition; Breastfeeding; Caregiving environment
	• Paternal Factors
	  o IQ; Education
	• Other Exposures
	  o Other prescribed medications; Alcohol; Tobacco; Other recreational drug use.
	• Medical and Family History
	  o Neurodevelopmental problems; Physical illness; Mental illness; Genetic disorder
	• Infant Factors
	  o Birthweight; Anthropometric data (length, head circumference etc); Sex; Prematurity; Physical illness; Physical injury; Early intervention

Exposure Details There was a clear theme in the comments that dose, timing and duration of the exposure were important considerations. To help us formalise this level of agreement to what extent do you agree that the following information should be considered in any analysis:	• Dose of exposure • Timing of exposure during pregnancy (e.g. onset in 1st, 2nd, 3rd trimester etc.) • Duration of exposure • Changes to medication during pregnancy/continuous measure of exposure • Route or mechanism of the drug • Patient compliance/adherence to prescribed dose • Continued collection of data throughout breastfeeding. • Exposure to maternal condition (e.g. seizures). • Biomarkers of the level of the medication present in the infant. • Co-medications
Recruitment (Case): To what extent to you agree that each of these statements would provide an optimal participant selection:	• Prospective recruitment of participants during pregnancy for longer term follow up • A sample that is representative of the relevant population (e.g. large sample, or through population-wide registers, multi-centre recruitment etc) • Pre-defined and standardised inclusion and exclusion criteria • Recruitment of participants for neurodevelopmental follow up from existing prospective medical or pregnancy registers (e.g. disease specific registers such as a register for women with epilepsy where initial recruitment was only focused on immediate birth outcomes)

Recruitment (Control/comparator) To what extent do you agree that the following provide one of the needed control/ comparator groups for medication exposure studies:	• The recruitment of participants who have been exposed to the same disease but a different medication • The recruitment of participants who have been exposed to the same disease but have not received prescribed medication. • The recruitment of participants who have not been exposed to the disease. • Do you agree that studies should recruit both comparator (e.g. same disease with different drug or no drug) AND control (e.g. no disease or drug) groups? • Control/comparators should be matched to exposed group on key maternal variables. • Control comparators should be recruited from the same clinic/community as the exposed group.
Exposure Data Ascertainment A number of methods for ascertaining exposure data were suggested by respondents. Please rank these in order of reliability, with 1 being the most reliable and x being acceptable but with limitations:	• Prospective concurrent self-report of medication use by women themselves • Collection of biomarker data (e.g. serum levels, umbilical cord blood) • Collection of medication exposure data from medical records • Collection of medication exposure data from pharmacy dispensing records. • A combined approach of collecting both self-report data from the mother and supplementary secondary data from medical records/ pharmacy dispensing records
Measurement of Neurodevelopment To what extent do you agree with the following statements regarding the optimal approach to outcome data collection:	• For certain outcome (e.g. infant development or child cognitive functioning) direct researcher led assessments are optimal (gold standard) • For other outcomes (e.g. child behaviour or social development) parent completed questionnaires are optimal • Assessments should be standardised across all included children • The assessment should be blinded to the exposure history of the child • Assessments should be administered by a small pool of researchers to reduce issues of inter-rater reliability.
Confounding Optimization Do you agree that the optimal way of selecting confounders for specific studies involves the following	• The development of a core set of pre-defined factors that have been discussed and agreed upon by experts in the field. • Although it may be possible to develop a core set of confounders, it remains important to consider and include additional confounders that may be of specific relevance to a given medication. • Any decisions regarding confounder or mediator selection should be made *a priori* and based on previous literature or scientific knowledge • Adequate consideration should be given to the distinction between *confounders* and *mediators,* to avoid masking potential and important indirect effects. • The use of novel techniques to detect previously unmeasured confounding variables (e.g. natural experiments, machine learning etc.) • Where the siblings are discordant for the medication exposure in question this addresses a number of important confounders
Attrition	• Attrition is an inherent characteristic within longitudinal cohort designs

Please indicate the extent to which you agree or disagree with the following statements:	• Systematic attrition (where one group or factor is more inclined to be lost to follow up) creates more of a challenge to the reliability of the results than a balanced attrition
	• What is an acceptable level of attrition in longitudinal studies, following patients for a number of years:
	  o 10%
	  o 20%
	  o 30%
	  o 40%
	• Predetermined statistical methods for dealing with unbalanced attrition should be in place.

Good reporting Please indicate the degree to which you agree or disagree with the following statements regarding good practice in reporting and interpreting results	• Absolute risk statistics and effect sizes (e.g. percentages, numbers within clinical ranges, mean differences etc) as well as relative risk or odds ratios results should be reported. • The context of the size of the effect or difference should be described with age related and national norms. • Standardised reporting of effect sizes Cohen's d would assist in comparing outcomes across studies and measures • Authors should provide context on how any effect might impact an individual's educational, occupational and relational attainment (where relevant).